

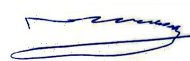
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ імені ІГОРЯ СІКОРСЬКОГО»

Факультет Електроніки
(повна назва інституту/факультету)

Кафедра акустичних та мультимедійних електронних систем
(повна назва кафедри)

«До захисту допущено»

В.о. завідувача кафедри

 С. А. Найда
(ініціали, прізвище)

« 01 » червня 20 20 р.


Дипломна робота
на здобуття ступеня бакалавра


зі спеціальності (спеціалізації) 171 - Електроніка
(код та назва спеціальності)

на тему: «Вплив кодеків на розбірливість мови в лініях зв'язку»

Виконав: студент 4 курсу, групи ДГ-г61-1
(шифр групи)


Ліщук Іван Олександрович
(прізвище, ім'я, по батькові) 
(підпис)

Керівник асистент Котвицький Ігор Валерійович
(посада, науковий ступінь, вчене звання, прізвище та ініціали) 
(підпис)

Консультант Якість кодованих сигналів, професор, д.т.н., Продеус А.М.
(назва розділу) (посада, вчене звання, науковий ступінь, прізвище, ініціали) 
(підпис)

Рецензент доцент кафедри ЕІ, к.т.н. Шуляк О.П.
(посада, науковий ступінь, вчене звання, науковий ступінь, прізвище та ініціали) 
(підпис)

Засвідчую, що у цій дипломній роботі
немає запозичень з праць інших авторів
без відповідних посилань.

Студент 
(підпис)

Київ – 2020 року

**Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»**

Інститут/факультет _____ Факультет електроніки _____
(повна назва)

Кафедра акустичних та мультимедійних електронних систем _____
(повна назва)

Рівень вищої освіти – перший (бакалаврський)

Спеціальність (спеціалізація) _____ 171 - Електроніка _____
(код і назва)

ЗАТВЕРДЖУЮ

В.о. завідувача кафедри

 С. А. Найда
(ініціали, прізвище)

«01» _____ червня 2020 р.

**ЗАВДАННЯ
на дипломний проект (роботу) студенту**

_____ Ліщук Іван Олександрович _____
(прізвище, ім'я, по батькові)

1. Тема проекту (роботи) «Вплив кодеків на розбірливість мови в лініях зв'язку» _____

керівник проекту (роботи) _____ асистент Котвицький І. В. _____,
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом по університету від «25» ____ 05 ____ 2020 р. № 1196-с



2. Строк подання студентом проекту (роботи) 14 червня 2020 р.

3. Вихідні дані до проекту (роботи) Наукові джерела з питань кодування та декодування сигналів, порівняння лінійного та нелінійного кодування, оцінювання якості спотворених сигналів _____

4. Зміст (дипломної роботи) пояснювальної записки (перелік завдань, які потрібно розробити) Огляд літературних джерел, класифікація кодеків, оцінювання якості аудіосигналів: суб'єктивне та об'єктивне, порівняння кодеків за критерієм якості сигналу

5. Перелік графічного (ілюстративного) матеріалу (із зазначенням обов'язкових креслеників, плакатів, презентацій тощо) 15 слайдів презентації PowerPoint

6. Консультанти розділів проекту (роботи)*

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Якість кодованих сигналів	професор, д.т.н., Продеус А.М.	 15.12.2019	 18.05.2020

7. Дата видачі завдання 01.12.2019

Календарний план

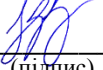
№ з/п	Назва етапів виконання дипломного проекту (роботи)	Строк виконання етапів проекту (роботи)	Примітка
1	Видача завдання	1.12.2019	Виконано
2	Аналіз літературних джерел	13.04 – 19.04	Виконано
3	Дослідження особливостей лінійного квантування	20.04 – 26.04	Виконано
4	Дослідження особливостей нелінійного квантування	27.04 – 03.05	Виконано
5	Закінчення роботи. Підведення підсумків. Оформлення роботи.	18.05 – 05.06	Виконано

Студент


(підпис)

Ліщук І.О.
(ініціали, прізвище)

Керівник проекту (роботи)


(підпис)

Котвицький І. В.
(ініціали, прізвище)

* Консультантом не може бути зазначено керівника дипломного проекту (роботи)

АНОТАЦІЯ

В роботі розглянуто питання оцінювання розбірливості мови в лініях зв'язку, що містять кодек G.711. Виконано класифікацію кодеків. Проведено моделюванням кодеку G.711 в середовищі Matlab. При цьому цифрові сигнали представлялися як числові масиви, отримані шляхом дискретизації та квантування. Оцінювання якості сигналів виконувалося суб'єктивним методом, шляхом порівняння на слух звучання неспотворених (на вході кодеку) та спотворених (на виході кодеку) сигналів.

ANNOTATION

The work considers the issue of language intelligibility assessment in communication lines. Detailed acquaintance with a variety of codecs. Introduction to G.711 codec simulation. The evaluation of signal quality was performed subjectively. Digital signals were given as numbers separated by time and amplitude. Separation by time is a process of sampling. By amplitude - quantization. Execution of G.711 codec modulation. Comparison of graphics and sound. The sound is listened through headphones.

ЗМІСТ

ВСТУП	7
РОЗДІЛ 1. АУДІОКОДЕКИ ТА ЇХ РІЗНОВИДИ.....	8
1.1. Класифікація кодеків	8
1.2. Аудіокодеки VoIP	16
1.3. Таблиця підсумків кодеків	22
1.4. Процес кодування в кодеку G.711	25
1.5. Моделювання квантування в Matlab	34
1.6. Кодек GSC: мовний кодек GSM	35
ВИСНОВКИ до розділу 1	40
РОЗДІЛ 2. ЯКІСТЬ КОДОВАНИХ СИГНАЛІВ	41
ВИСНОВКИ до розділу 2	46
ВИСНОВКИ.....	47
ПЕРЕЛІК ПОСИЛАНЬ	48
Додаток А. Програма моделювання кодеку G.711 в середовищі Matlab	50

ВСТУП

Аудіо кодеки – це результат прогресу в галузі цифрової телефонії, який дозволив здійснювати голосові та відео дзвінки з одного кінця планети на інший, перетворюючи аналогові аудіосигнали в цифрові дані, які можуть передаватися через Інтернет за конкретною IP-адресою.

Однак це стало можливим лише завдяки звуковим кодекам (аббревіатура для кодера-декодера), які виконують це завдання перетворення, беручи вибірки (миттєві значення) аудіосигналу тисячі разів за секунду, а потім перетворюючи їх у цифрову інформацію. Наприклад, у випадку кодека G.711 береться 64000 вибірок в секунду. Дуже корисним способом вимірювання якості кодека є використання міри суб'єктивної якості сигналу MOS (середня оцінка думки). Ця міра дозволяє суб'єктивно, шляхом прослуховування людини, оцінити якість звуку після всього процесу стиснення, передачі та декомпресії. Значення MOS визначається оцінками 1 і 5, де 1 – поганий, а 5 – відмінний.

Метою даної роботи є експериментальне порівняння якості кодеків шляхом суб'єктивної та об'єктивної оцінки якості звукових сигналів, які піддавалися компресії та декомпресії із використанням процедури кодування.

Доцільність такого порівняння зумовлена тим, що стиснення кодеками дозволяє суттєво зменшити об'єм даних, що передаються лініями зв'язу, проте відбувається це ціною зменшення якості сигналу.

Результати проведеної роботи будуть корисними для тих хто працює чи працюватиме в сфері комунікаційної акустики та, зокрема, в сфері IP-телефонії.

РОЗДІЛ 1

АУДІОКОДЕКИ ТА ЇХ РІЗНОВИДИ

Існує два види кодеків: аудіокодеки та відеокодеки. В даному розділі ми будемо розглядати різновиди аудіокодеків.

1.1. Класифікація кодеків

Кодеки діляться на два типи: перший – **програмний кодек**, який є комп'ютерною програмою, та другий – це **програмно-апаратний кодек** («апаратний»), він в свою чергу позначає якусь окрему конкретну мікросхему. Програмні та апаратні кодеки мають багато спільного. Поряд з тими ж параметрами якості приміщення, програмні кодеки можуть пропонувати функції безпеки, такі як захист паролем і шифрування від кінця до кінця. Однак важливо перевірити особливості безпеки будь-якої програмної платформи, перш ніж завершити вибір. Наприклад, у специфікаціях Skype for Business вказано, що їх комунікація захищена 256-бітним шифруванням TLS AES, яке відповідає стандартам Міністерства оборони (DOD) для шифрованих комунікацій. Розглянемо ці два типи детально.

Архівація переговорів можлива як з апаратними, так і з програмними кодеками, але, як правило, є ріднішими для програмних кодеків, що мають коріння в хмарній інфраструктурі. Для запису переговорів з апаратними кодеками може знадобитися додаткова інфраструктура. Обидва варіанти можуть поєднувати візуальні відеоконференції та обмін вмістом. М'які кодеки можуть додавати спільні функції, які дозволяють всім учасникам одночасно розміщувати вміст, тоді як апаратні кодеки, як правило, підтримують лише один хост, який надає спільний вміст одночасно.

Апаратні кодеки хороші для підключення приміщень «точка-точка», якщо в них обох встановлений однаковий апаратний кодек виробника. Але в міру

того, як робоча сила стає більш мобільною, це менш ймовірно. Для підключення декількох апаратних кодеків, підключення кодеків різних виробників та переходу на мобільних клієнтів можна використовувати послугу мосту або віртуальну кімнату для нарад (VMR). Але послуга мосту або VMR - це ще одна додаткова вартість і може не піддаватись імпровізованим зустрічам. Для порівняння, оскільки програмні кодеки - це локальні хмарні технології, вони не потребують обслуговування мостів, а додавання кількох відвідувачів на ходу є досить тривіальним.

Контроль конференц-залу м'яких клієнтів є складнішим, ніж з апаратними кодеками. Оскільки кожен програмний кодек - це інша програма, що працює на ПК або мобільному пристрої, не існує фізичного інтерфейсу, який можна підключити до кімнатного контролера. Компроміс полягає в тому, що м'які кодеки підтримують безліч різних конференц-програм без необхідності обслуговування мостів. Досвід роботи в конференц-залі раніше був центральним у відеоконференціях, оскільки це був єдиний варіант, але оскільки мобільний перетворюється на домінуючий випадок використання, цей компроміс стає більш прийнятним. Ви повинні оцінити, що цінує ваша організація, і вирішити, чи віддаєте перевагу простоті керування в кімнаті або гнучкістю для легкого підключення до будь-якої веб-платформи. Типовими клієнтами, які вибирають апаратні кодеки, є юридичні фірми, фінансові компанії.

Як бізнес-рішення, це зводиться до беззбиткового порівняння власності на обладнання та моделі вартості підписки на програмне забезпечення, а також до оцінки інших перелічених вище міркувань. Skype for Business є поширеним у корпоративному світі. Якщо ваша організація вже використовує Skype for Business для веб-конференцій, розгляньте можливість включення веб-конференцій в конференц-залі за допомогою платформи Skype. Якщо ви використовуєте лише Skype як клієнт чату і вам потрібно вибрати клієнта для чату, вам потрібно буде порівняти функції з вартістю. Але ми не схвалюємо

будь-яку платформу конференцій або системну архітектуру для підходу до конференцій. Ваша організація повинна зважити компроміси та визначити, що найкраще підходить для вашої культури на робочому місці та колективу технічної підтримки.

Аудіокодек на програмному рівні є спеціалізованою комп'ютерною програмою, кодеком, який стискає (виконує компресію) або розтискає (виконує декомпресію) цифрові звукові дані відповідно до файловим звуковим форматом або потоковим звуковим форматом. Завдяки стиску зменшується обсяг простору, необхідного для зберігання звукової інформації, а також можливо знизити смугу пропускання каналу, по якому передаються аудіодані. Більшість аудіокодеків здійснені як програмні бібліотеки. Розглянемо ці бібліотеки:

- **AudioTX Communicator**

Використовується для зв'язку студії з віддаленим передавачем. Передача аудіо даних через ISDN, мережі, DSL, ADSL та Internet. Кодек AudioTX Communicator сумісний з виробниками ISDN кодеків - CDQ Prima, Telos Zephyr, Glensound, Dialog, Philips. Кодування в MPEG 2 MPEG 3 G.722 G.711. Програма працює під Windows 98 / NT / 2000 / Me / XP. Використання глобальних мереж, для передачі звуку, що надійно і економічно вигідно. Широкомовна передача звуку якості по ISDN, IP мережами і Internet, використовуються MPEG2, MPEG3, G.722 і G.711 або лінійний (нестислий) звук.

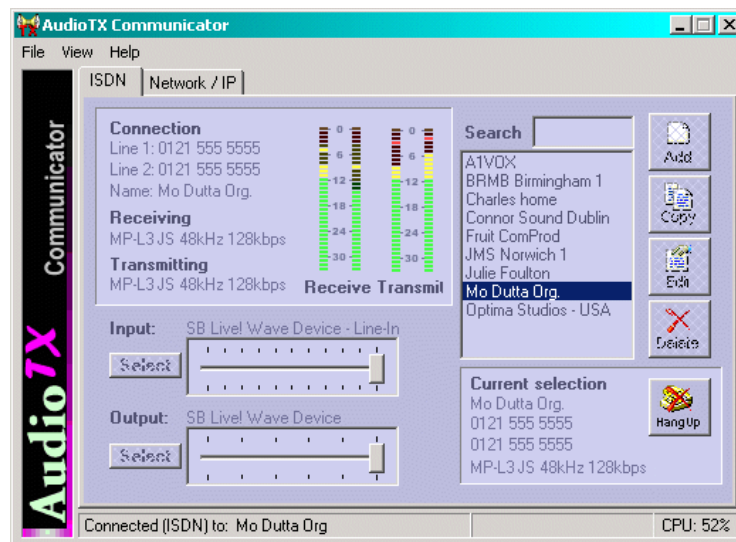


Рис.1.1. Інтерфейс програми «AudioTX Communicator»

- **Loudness Analyzer**

Це професійний додаток для контролю рівня гучності ефіру. Він аналізує і нормалізує рівень гучності медіафайлів відповідно до стандарту EBU R 128. На сьогоднішній день Loudness Analyzer - це єдиний вітчизняний програмний продукт, сертифікований за системою добровільної сертифікації програмного забезпечення для засобів вимірювань.

Открыть файлы...

4 / 4

Видео файл	Количество дорожек	Длительность					
> 8522504	8	00:00:10					
> 8522505	8	00:00:07					
< 8522508	8	00:00:05					
Аудиодорожка	Количество каналов	Длительность	Максимальный уровень пиковых значений, dBTP	Максимальный уровень интегрированной громкости, LUFS	Максимальный уровень кратковременной громкости, LUFS	Интегральная громкость, LUFS	Динамический диапазон, LU
Аудиодорожка 1	1	00:00:07	-47.3	-61.8	-61.9	-61.9	0.0
Аудиодорожка 2	1	00:00:07	-7.0	-25.1	-30.4	-31.5	3.6
Аудиодорожка 3	1	00:00:07	-7.0	-23.2	-29.2	-17.6	6.3
Аудиодорожка 4	1	00:00:07	-7.7	-23.3	-29.3	-17.3	5.5
Аудиодорожка 5	1	00:00:07	-70.0	-70.0	-70.0	-70.0	0.0
Аудиодорожка 6	1	00:00:07	-70.0	-70.0	-61.9	-70.0	0.0
Аудиодорожка 7	1	00:00:07	-70.0	-70.0	-70.0	-70.0	0.0
Аудиодорожка 8	1	00:00:07	-70.0	-70.0	-70.0	-70.0	0.0

Нормализовать

Сохранить

Loudness Analyzer

Рис.1.2. Інтерфейс додатку «Loudness Analyzer»

- **Stereo Tool**

Це програмний аудіопроцесор, який забезпечує видатну якість звуку і володіє багатьма унікальними функціями. Він використовується більш ніж 1500 FM-станціями, починаючи від невеликих місцевих станцій, до станцій потужністю 50-100 кВт і загальнонаціональних мереж з десятками передавачів, тисячами потокових станцій і багатьма станціями DAB +, HD, AM і TV. Його можна використовувати як для оперативної, так і для файлової обробки.

Програма Stereo Tool представляє з себе панель управління звуком з безліччю налаштувань і регуляторів для налаштування звуку. Крім того, в даному плагіні є встановлені пресети для швидкого налаштування якості звучання. Основне призначення Stereo Tool - посилювати навіть слабкий сигнал при прослуховуванні FM-радіо і MP3-треків в низькій якості. Крім того, плагін дозволяє конвертувати стерео в моно без втрати якості звучання.

Програма може бути використана і як звуковий плагін для популярного медіаплеєра Winamp.

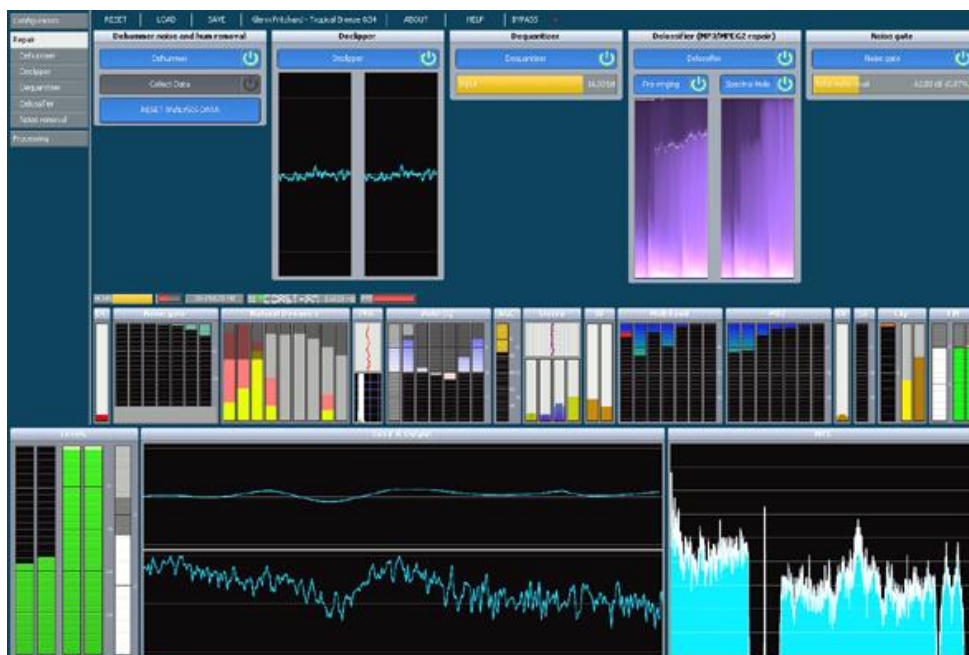


Рис.1.3. Інтерфейс програми «Stereo Tool»

- **RadioLogger**

Дана програма, призначена для запису ефіру радіостанції на жорсткому диску комп'ютера. Запис може проводитися 24 години на добу або в задані тимчасові інтервали. Використовуються найсучасніші методи стиснення звуку,

що дозволяє зберігати записи в хорошій якості і при цьому не займати багато місця на диску.

Переваги:

- автоматичний старт / стоп запису в заданий час (планувальник);
- кодування в mp3, ogg або іншим кодеком "на льоту" (без проміжної записи в WAV);
- низьке завантаження процесора;
- ведення історії записів (архів);
- автоматичне видалення старих записів;
- безліч налаштувань;
- може застосовуватися для записи з різних джерел, наприклад, мікрофона;
- автоматична розбивка записи на частини;
- багатомовний інтерфейс.

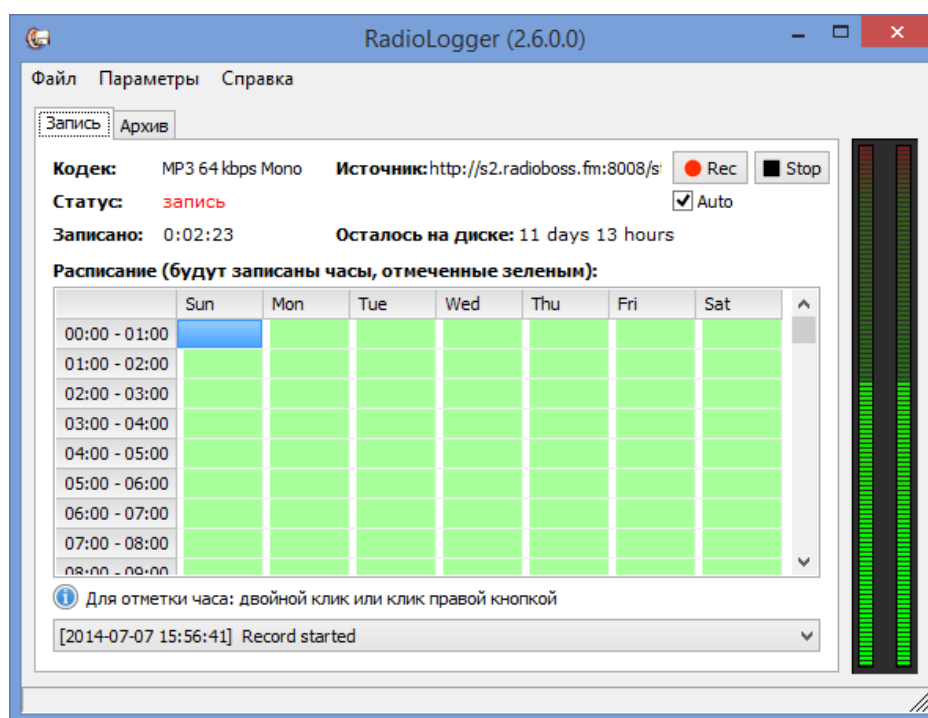


Рис.1.4. Інтерфейс програми «RadioLogger»

- **Махх**

Програма є інструментом для остаточного обмеження аудіо в продакшн середовищі. Працює в інтерактивному режимі (ручному) і має пакетну обробку файлів режимі з використанням нового двосмугового (MBL2) процесора або

дев'яти груп процесорів, заснованих на MBL9 від Sonos 4. Обидва обробляють спотворення і маскують їх остаточно кліпером (обмежувачем).

Махх використовує два проходи обробки. Перший прохід нормалізує рівень вхідного сигналу. Це дозволяє уникнути використання AGC, крім остаточної обробки.

Архітектура двох процесорів аналогічна. Додатковий двохдіапазонний AGC, далі опціонально багатосмуговий і широкосмуговий обмежувач і, нарешті кінцевий кліпер (обмежувач). В ручному інтерактивному режимі є прості у використанні елементи управління для налаштування процесорів. Пресети можуть бути збережені або завантажені. MBL2 є більш прозорим, ніж дев'ять груп процесора (MBL9), але звучить не так голосно.

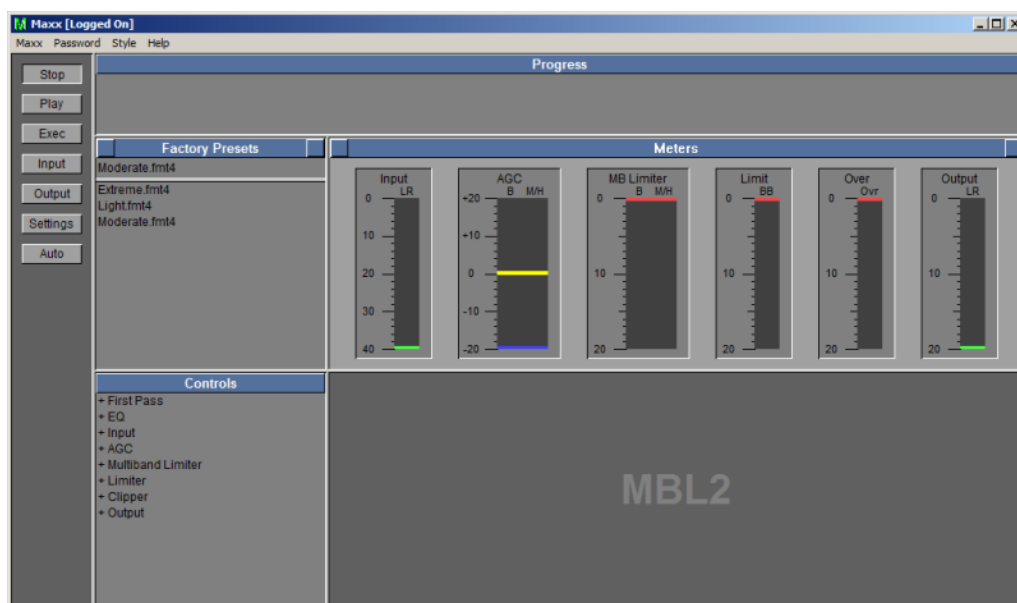


Рис.1.5. Інтерфейс програми «Махх»

- **Sonos 4**

Sonos 4 є повною звуковою обробкою і компонується (STL) з системами для радіомовлення. Тепер стало можливо, контролювати ефір, відразу двома бібліотеками MBL5 & MBL8.

У програмного забезпечення є такі головні особливості:

- Дозволяє одному комп'ютеру запустити до 8 процесорів одночасно.

- Підтримка інтерфейсу Windows і інтерфейсу ASIO. Частота дискретизації до 192 кГц і 32 біт.
- Має 5 процесорів в групі і 9 груп процесорів. Параметри настройки процесора повністю коректуючого зі зручним збереженням і завантаженням файлів. Є планування по днях тижня (Daypart). Всі чотири процесори спільно використовують загальні частини, такі як три стадії параметричного зрівнювання, розширення стерео, зміна фази, фільтрація для FM і AM, з перемиканням Pre-Emphasis для FM. Архітектура процесора схожа на кожному процесорі. Одна група Gated AGC включає в себе багатосмуговий лімітер, можливе включення перед ним багатосмугового компресора. Додатково одна група обмеження Gated може застосовуватися до фінального обмежувача (clipper). Pre-Emphasis для FM застосовується частково до багатосмугової секції і частково після фінального обмежувача.
- В межах кожного процесора є 'запобіжна' функція, яка переключиться, до резервної подачі аудіо, якщо основний сигнал дасть збій. Резерв може бути просто плеєром файлу, що запускає потрібний файл з музикою.
- Мультиплексний стерео генератор і кодер RDS використовується в FM радіомовленні. Потребує звуковій платі з адекватною смугою пропускання 192 кГц, щоб передати мультиплексний сигнал.
- Є зв'язок з віддаленим комп'ютером по локальній мережі або інтернет (функція STL). Можна послати сигнал кільком клієнтам. Посилання довільна і двонаправлена. Кодер використовує Ogg/Vorbis FM якості, 64kb/s.
- Функція реєстрації аудіо.



Рис.1.6. Інтерфейс програми «Sonos 4»

1.2. Аудіокодеки VoIP

Як і у випадку оцифрування мови для передачі телефонії, кодеки працюють за принципом вибірки, квантування та кодування. Це дозволяє досягти оптимальної якості мови для передачі аналогового мовного сигналу через цифрові системи.

У мережі з комутацією пакетів завжди прагнуть зберегти пропускну здатність. Ось чому голосові дані зазвичай все ще стискаються. Однак зі збільшенням стиснення якість мови знижується, збільшується час декомпресії та необхідна обчислювальна потужність. Однак це не завжди стосується. Більшість кодеків використовують властивості людської мови стискатися з якомога меншими втратами.

Кодеки оцифрування мовлення повинні відповідати деяким умовам. Завдяки структурі Інтернету (передача та комутація, орієнтована на пакет), вони повинні мати можливість справлятися з втратами пакетів (до 5%)

та компенсувати різниці в режимі виконання окремих пакетів та сортувати їх у правильному порядку (виправлення помилок уперед та буферизація тремтіння). Це означає, що втрати пакетів та коливання часу виконання не повинні впливати на якість голосу.

MOS – середній показник для VoIP

MOS або середній показник думок дає тестування VoIP через числове значення як вказівку на сприйману якість одержуваного голосу після передачі та стискання за допомогою кодеків. Це вимірювання є результатом базових мережевих атрибутів, які впливають на потік даних і корисні для прогнозування якості дзвінків, і є хорошим інструментом тестування VoIP для визначення питань, які можуть впливати на якість VoIP та ваші розмови.

Використання MOS балів для оцінки вашої VoIP-послуги.

Тестування якості VoIP стало простішим, а послуги значно покращилися за останні кілька років через те, що провайдери стають більш надійними та Інтернет-провайдерами, які пропонують кращі зв'язки. Це підвищення якості послуг допомогло збільшити кількість абонентів VoIP, але періодично виникають проблеми, які впливають на якість голосу, і можливість тестування вашого VoIP та виявлення цих випадків може бути корисною для їх вирішення. Встановлення показника для вимірювання змін або погіршення якості голосового / VoIP-з'єднання після тестування може допомогти визначити проблеми. VoIP-дзвінки часто знаходяться в діапазоні від 3,5 до 4,2 MOS. Наведена нижче таблиця може бути використана в якості посібника для тестування MOS та гарного порівняння якості голосу.

Значення порівняння MOS

Дуже задоволений	4,3-5,0
Задоволений	4,0-4,3
Деякі користувачі задоволені	3,6-4,0
Багато користувачів незадоволені	3.1-3.6

Майже всі користувачі незадоволені	2.6-3.1
Не рекомендовано	1,0-2,6

Таблиця 1. Значення порівняння MOS

Що впливає на тест MOS?

По-перше, важливо розуміти, що MOS - середня оцінка думки - є відносною шкалою і ґрунтується на багатьох факторах, які можуть впливати на якість голосу. Вимірювання VoIP збирають після тестування однобічної затримки або затримки з'єднання, втрати пакетів з метрикою, що включає кількість послідовних втрачених пакетів, і кількість тремтіння (різниця у часі, на який потрібно приймати пакети). Потім обчислення враховують коефіцієнт **R**, який може бути використаний для оцінки показника MOS.

Затримка розповсюдження - це час, необхідний для цифрового сигналу, що проходить в кінці по всій мережі. Чим більше відстань, тим більша затримка розповсюдження. Крім того, дані мають проходити через мережеві маршрутизатори, комутатори та інші пристрої, такі як міжмережеві стіни, кожен додаючи власну (транспортну) затримку. **Затримка пакетизації** - це час, необхідний для оцифрування сигналу для кодека, який використовується для надсилання через Інтернет та декодування його в далекому кінці. Більш стислий кодек, як G.729, має більшу затримку пакетизації, ніж нестиснений кодек, як кодек G.711. **Jitter буфер** - це затримка, введена пристроєм АТА для утримання однієї або декількох дейтаграм, для компенсації варіацій часу прибуття. У VoIP буфер тремтіння - це область, де голосові пакети можна збирати, зберігати та потім надсилати процесору з більш рівномірно розташованими інтервалами. Коливання часу прибуття пакетів (тремтіння) зазвичай є результатом перевантаженості мережі або зміни маршруту. Буфер тремтіння, який розташований на приймальному кінцевому пристрої голосового з'єднання, навмисно затримує пакети, щоб отриманий голос був представлений правильно з меншими спотвореннями.

Наступні елементи можуть впливати на якість дзвінків:

- Пропускна здатність
- Кодек у використанні.
- Обладнання
- Джиттер
- Затримка
- Втрата пакетів

Протоколи VoIP

Протягом багатьох років існував чіткий поділ між двома типами мереж:

- **голосові мережі**, засновані на комутації ланцюгів, ланцюг зарезервований і маршрутизація під час зв'язку завжди проводиться одним і тим же шляхом. Приклад: загальнодоступна комутаційна телефонна мережа (PSTN);
- **мережі передачі даних** на основі комутації пакетів, інформація поділяється на пакети і кожен пакет може проходити різним шляхом. Приклад: Інтернет Для того, щоб надсилати інформацію через мережі передачі даних (Інтернет), що базуються на комутації пакетів, необхідно прийняти протоколи, відповідальні за передачу та відновлення інформації.

Проблема технології комутації схем полягає в тому, що вона вимагає великої кількості пропускної здатності для кожного дзвінка. Схема використовується не ефективно, тому що той самий канал використовується під час дзвінка, і більшість телефонних дзвінків мають багато мовчання.

Мережі даних, навпаки, передають інформацію лише тоді, коли це необхідно, ефективно використовуючи пропускну здатність. Затримка, варіації затримки доставки пакетів та втрата пакетів не повинні бути недоліком, оскільки кінцева система має процедури відновлення вихідної інформації. Однак передача голосу та відео дуже чутливі до цих

параметрів. Тому необхідні мережі та протоколи з високим рівнем якості QoS (якість обслуговування) .

Voice over IP (VoIP) визначає необхідні системи маршрутизації та протоколи для передачі голосових розмов через Інтернет. Інтернет - мережа комутації пакетів на основі протоколу TCP/IP.

На даний момент, в основному, є дві архітектури VoIP для передачі голосу через Інтернет, які дуже використовуються:

- SIP (протокол ініціації сесії).
- H.323.

SIP Архітектура

Протокол ініціації сеансу (SIP) - це протокол, розроблений робочою групою IETF MMUSIC та запропонований стандарт для ініціювання, зміни та завершення інтерактивного сеансу користувача, який включає мультимедійні елементи, такі як відео, голос, обмін миттєвими повідомленнями, онлайн-ігри та віртуальна реальність . Спочатку він був опублікований у 1996 році як RFC 2543, тепер застарілий, завдяки публікації нового RFC 3261 у 2002 році.

SIP може встановлювати сеанси для функцій, таких як аудіо/відеоконференції, інтерактивні ігри та переадресація дзвінків, що розгортаються через IP-мережі, що дозволяє постачальникам послуг інтегрувати основні послуги IP-телефонії з веб-послугами, електронною поштою та послугами чату. Приклад - SIP-тренкінг від MegaPath.

Основна мета SIP - це зв'язок між мультимедійними пристроями. SIP робить можливим спілкування завдяки двом протоколам: RTP / RTCP та SDP.

Протокол RTP використовується для транспортування голосових даних у режимі реального часу (те саме, що H.323 протокол), тоді як протокол SDP використовується для узгодження можливостей учасника, типу кодифікації тощо.

SIP розроблений відповідно до Інтернет-моделі. Це протокол сигналізації, орієнтований на кінець, який означає, що вся логіка зберігається в кінцевих пристроях (крім маршрутизації SIP-повідомлень). Стан також зберігається лише в кінцевих пристроях, не існує єдиної точки відмови, і мережі добре розроблені таким чином. Ціна, яку нам доведеться платити за дистрибутивність та масштабованість, - це більші накладні витрати, спричинені тим, що повідомлення надсилаються в кінці.

Тому SIP - це протокол управління на рівні додатків, протокол сигналізації для Інтернет-телефонії. SIP може встановлювати сеанси для функцій, таких як аудіо/відеоконференції, інтерактивні ігри та переадресація дзвінків, що розгортаються через IP-мережі таким чином, дозволяючи постачальникам послуг інтегрувати основні послуги IP-телефонії з веб-послугами, електронною поштою та послугами чату. Він заснований на запитах та відповідях і повторно використовує багато концепцій попередніх стандартів, таких як HTTP та SMTP.

H.323 протокол

H.323 був розроблений з основною ціллю: забезпечити телеконференцію з голосовими, відео та інформаційними можливостями в мережах комутації пакетів.

Безперервні дослідження та розробки H.323 дотримуються тієї ж мети. Крім того, H.323 та конвергенція голосу, відео та даних дозволяють постачальникам послуг надавати нові можливості для користувачів та покращувати продуктивність для користувача, зменшуючи витрати.

Стандарт був розроблений спеціально для досягнення наступних цілей:

- для того, щоб бути на основі існуючих стандартів, включаючи H.320, RTP і Q.931;
- щоб додати деякі з переваг, які пакетна комутація мережа пропонує транспортувати реальні дані про час;

- для вирішення проблеми даних у режимі реального часу в мережах комутації пакетів.

Дизайнери H.323 знають, що вимоги до комунікації відрізняються від місця до місця, між користувачами та між компаніями, і очевидно, що вимоги майбутніх додатків також змінюються. Отже, дизайнери H.323 визначили це таким чином, що компанії, що виробляють обладнання, можуть додавати до протоколу власні специфікації та визначати інші структури та стандарти, які дозволяють пристроям придбати нові функції або потужності.

1.3. Таблиця підсумків кодеків

Передача голосу є аналогічною, тоді як мережа передачі даних є цифровою. **Процес вибірки аналогічних хвиль в цифрову інформацію проводиться кодером-декодером (CODEC).** Існує багато стандартів для вибірки аналогового голосового сигналу в цифровий. Процес часто буває досить складним. Більшість перетворень використовують імпульсну модуляцію коду (PCM) або варіації.

Крім того, CODEC поштовує послідовність даних, а іноді забезпечує ехо-скасування. Стиснення форми хвилі може зберегти пропускну здатність. Це особливо цікаво при низькошвидкісних з'єднаннях, щоб ви могли мати більше VoIP-з'єднань одночасно. Ще одним способом збереження пропускну здатності є використання придушення тиші. Мета - не надсилати пакунки, коли в розмовах немає голосу.

Далі наведена таблиця з найбільш відомими кодеками, які використовуються:

- Швидкість передачі бітів - Швидкість, з якою біти передаються по шляху зв'язку. Зазвичай виражається в кілобітах за секунду (Кбіт/с).
- Частота дискретизації - кількість проб, взятих за секунду при оцифровці звуку. Якість цифрового відтворення покращується зі збільшенням кількості взятих зразків за секунду.

- Розмір кадру - час між надісланими пакетами.
- MOS - (Середня оцінка думки). Це суб'єктивний показник якості звуку від 1 до 5.

Номер	Стандартний вид	Опис	Швидкість передачі (кб / с)	Швидкість вибірки (кГц)	Розмір кадру (мс)	Зауваження	MOS (середня оцінка думки)
G.711	MCE-T	Імпульсна модуляція коду (PCM)	64	8	Проба	U-law (США, Японія) та A-law (Європа)	4.1
G.711.1	MCE-T	Імпульсна модуляція коду (PCM)	80-96 Кбіт / с	8	Проба	Удосконалення та G.711 для забезпечення пропускну здатності аудіо від 50 Гц до 7 кГц	4.1
G.721	MCE-T	Адаптивна диференціальна імпульсна кодова модуляція (ADPCM)	32	8	Проба	Тепер описано в G.726; застарілий.	
G.722	MCE-T	7 кГц аудіокодування в межах 64 кбіт / с	64	16	Проба	Піддіапазон-кодек, який розділяє смугу 16 кГц на два піддіапазони, кожен кодований за допомогою ADPCM	
G.722.1	MCE-T	Кодування при 24 і 32 кбіт / с для вільної роботи в системах з низькими втратами кадру	24/32	16	20		
G.722.2 AMR-WB	MCE-T	Адаптивний широкосмуговий широкосмуговий кодек (AMR-WB)	23,85 / 23,05 / 19,85 / 18,25 / 15,85 / 14,25 / 12,65 / 8,85 / 6,6	16	20	В основному використовується для стиснення мовлення в мобільній телефонії 3-го покоління	
G.723	MCE-T	Розширення Рекомендації G.721 адаптивна диференціальна імпульсна кодова модуляція до 24 та 40 кбіт / с для застосування обладнання для множення цифрових схем	24/40	8	Проба	Витіснив G.726; застарілий. Це зовсім інший кодек, ніж G.723.1	3.65
G.723.1	MCE-T	Двошвидкісний мовний кодер для передачі мультимедійних комунікацій зі швидкістю 5,3 та 6,3 кбіт/с	5.6 / 6.3	8	30	Частина відеоконференцій H.324. Він кодує мовленнєві або інші звукові сигнали в кадрах, використовуючи лінійний прогнозовий аналіз за допомогою синтезу. Сигналом збудження для високошвидкісного кодера є багатоімпульсне максимальне вірогідне квантування (MP-MLQ), а для кодера	3.8-3.9

						низької швидкості - лінійне прогнозування, збуджене алгебраїчним кодом (ACELP).	
G.726	MCE-T	40, 32, 24, 16 кбіт / с адаптивна диференціальна імпульсна кодова модуляція (ADPCM)	16/24/32/40	8	Проба	ADPCM; замінює G.721 та G.723.	3.85
G.727	MCE-T	5-, 4-, 3- та 2-бітний / зразок вбудованої адаптивного диференціального імпульсного кодуляційного модуляції (ADPCM)			Проба	ADPCM. Пов'язане з G.726	
G.728	MCE-T	Кодування мови при 16 кбіт / с за допомогою коду з низькою затримкою збуджує лінійне прогнозування	16	8	2.5	CELP	3.61
G.729	MCE-T	Кодування мови при швидкості 8 кбіт / с за допомогою лінійного прогнозування, збудженого алгебраїчно-кодовим кодом (CS-ACELP)	8	8	10	Мала затримка (15 мс)	3.92
G.729.1	MCE-T	Кодування мови при швидкості 8 кбіт / с за допомогою лінійного прогнозування, збудженого алгебраїчно-кодовим кодом (CS-ACELP)	8/12/14/16 / 18/20/22/24 / 26/28/30/32	8	10	Удосконалення та G.711 для забезпечення пропускну здатності аудіо від 50 Гц до 7 кГц	
GSM 06.10	ETSI	Регулярне імпульсне збудження LongTerm Predictor (RPE-LTP)	13	8	22.5	Використовується для стільникової телефонії GSM.	
LPC10	Уряд США	Лінійно-прогнозний кодек	2.4	8	22.5	10 коефіцієнтів.	
Speex			8, 16, 3	2.15-24.6 (NB) 4-44.2 (WB)	30 (NB) 34 (WB)		
iLBC			8	13.3	30		
DoD CELP	Американський міністерство оборони (DoD) Уряд США		4.8		30		
EVRC	3GPP2	Підвищена змінна швидкість CODEC	9,6 / 4,8 / 1,2	8	20	Ось і покращує CDMA	
DVI	Інтерактивна мультимедійна асоціація (IMA)	DVI4 використовує адаптивну дельта-імпульсну модуляцію коду (ADPCM)	32	Змінна	Проба		
L16		Нестиснені зразки аудіоданих	128	Змінна	Проба		
SILK	Skype		Від 6 до 40	Змінна	20	Кодек гармонії пов'язаний з SILK	

Таблиця 2. Таблиця кодеків

G711 має дві версії під назвою U-law (США, Японія) та A-law (Європа). μ -закон відповідає стандарту T1, який використовується в Північній Америці та Японії. А-закон стосується стандарту E1, який використовується в іншому світі. Різниця полягає в способі вибірки аналогового сигналу. В обох схемах сигнал не дискретизується лінійно, а логарифмічно.

1.4. Процес кодування в кодеку G.711

Голосові та аудіосигнали аналогічні, тоді як мережа передачі даних є цифровою. Перетворення аналогового сигналу в цифровий здійснюється Аналого-цифровим перетворювачем (АЦП).

Цей процес аналого-цифрового перетворювача або модуляції імпульсного коду (PCM) виконується в три етапи:

- вибірки;
- квантування;
- кодифікація.

Вибірка

Вибірка - це процес кодування аналогового сигналу в цифровій формі шляхом зчитування (вибірки) його рівня з точно розташованими інтервалами у часі. Отримані значення називаються вибірками.

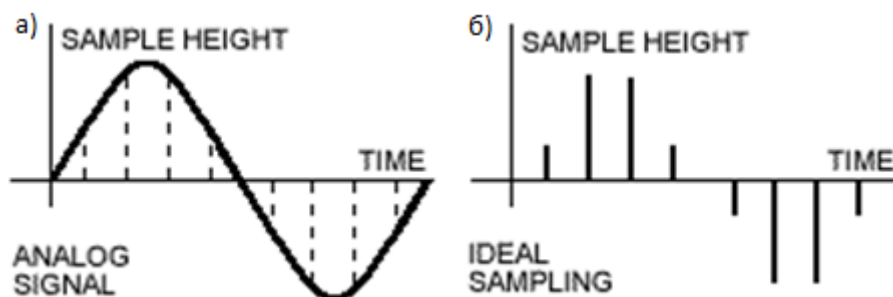


Рис.1.7. а) Аналоговий сигнал; б) Ідеальна вибірка

Відбір проб зазвичай відбувається з однаково відокремленими інтервалами; цей інтервал називається інтервалом вибірки. Зворотний інтервал вибірки називається частотою вибірки або частотою вибірки. Одиниця швидкості вибірки – Гц.

Умова, яка повинна відповідати частоті вибірки, задана теоремою вибірки: "Вказується, що смуговий сигнал з обмеженими частотними компонентами вище певної частоти відсічення однозначно визначається його дискретними значеннями в однаково розташованих точках, за умови, що ці зразки взяті зі швидкістю дискретизації, що дорівнює або перевищує подвійну частоту відключення." Відповідно до теореми вибірки, телефонні аудіосигнали (частотою від 300 Гц до 3400 Гц) повинні бути відібрані на частоті, рівній або більше 6800 Гц (2×3400).

Насправді ми зазвичай приймаємо частоту вибірки або частоту вибірки при 8000 Герц. Отже, беруть 8000 проб в секунду, які відповідають однаково розділеним інтервалам:

$$T = 1/8000 = 0,000125 \text{ сек.} = 125 \text{ мкс} \quad (1.1)$$

Тому дві послідовні вибірки одного і того ж сигналу відокремлюються 125 мкс, що називається інтервалом вибірки.

Квантування

Квантування - це процес переходу від континуальної множини значень сигналу до дискретного множини, обсяг якого дорівнює кількості рівнів квантування.

Рівномірне квантування

Рівномірне (лінійне) квантування. Розглянемо рис.1, на якому зображені L рівнів квантування аналогового сигналу з повним діапазоном напруг, рівним

$V_{pp} = V_p - (-V_p) = 2V_p$ вольт. Квантованні імпульси можуть мати позитивні і негативні-тільні значення. Крок між рівнями квантування, званий інтервалом квантування, становить q вольт. Якщо рівні квантування рівномірно розподілені по всьому діапазону, пристрій квантування іменується рівномірним, або лінійним. Кожне дискретне значення аналогового сигналу апроксимується квантованим імпульсом: апроксимація дає помилку, що не перевищує $q/2$ в позитивному напрямлені або $-q/2$ в негативному. Таким чином, погіршення сигналу внаслідок квантування обмежена половиною квантового інтервалу, $\pm q/2$ вольт.

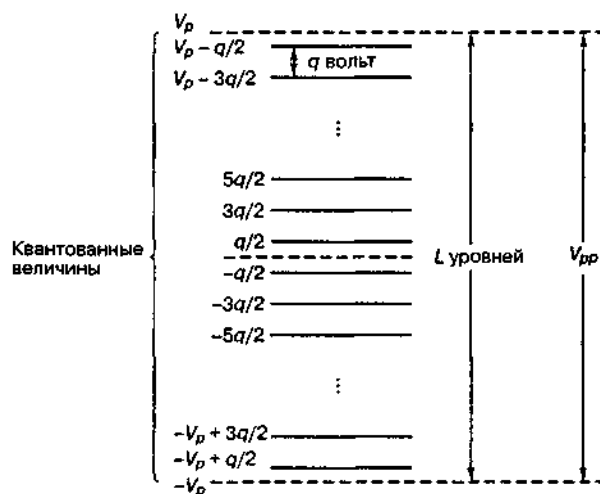


Рис.1.8. Рівні квантування

Хорошим критерієм якості рівномірного пристрою квантування є його дисперсія (середній квадрат помилки $e_i = \tilde{x}_i - x_i$ при нульовому середньому, де \tilde{x}_i - квантований віддік, x_i - квантуючий віддік). Якщо вважати, що помилка квантування e рівномірно розподілена в межах інтервала квантування шириною q (т. е. Приймає всі можливі значення з однаковою ймовірністю), то дисперсія помилок для пристрою квантування становить

$$\sigma^2 = \int_{-q/2}^{q/2} e^2 p(e) de = \int_{-q/2}^{q/2} e^2 \frac{1}{q} de = \frac{q^2}{12}$$

(1.2)

де p ($p(e) = 1/q$ при $-q/2 < q < q/2$ і 0 для $dp.q$ - щільність ймовірностей помилки квантування. Дисперсія σ^2 відповідає середній потужності шуму квантування. Пікову потужність аналогового сигналу (нормовану на 1 Ом) можна виразити як:

$$V_p^2 = \left(\frac{V_{pp}}{2} \right)^2 = \left(\frac{q(L-1)}{2} \right)^2 \approx \frac{q^2 L^2}{2} \quad (1.3)$$

де L - число рівнів квантування. Об'єднання виразів (1) і (2) дає відношення пікової потужності сигналу до середньої потужності квантового шуму $(S/N)_q$:

$$(S/N)_q = \frac{q^2 L^2 / 4}{q^2 / 12} = 3L^2, \quad (1.4)$$

де $N = \sigma^2$ - середня потужність шуму квантування. Очевидно, що відношення $(S/N)_q$ квадратично зростає з числом рівнів квантування. У межі ($L \rightarrow \infty$) сигнал стає аналоговим (нескінченне число рівнів квантування і нульовий шум квантування). Відзначимо, що для випадкових сигналів в параметр входить не максимальна, а середня потужність сигналу. У цьому випадку для отримання середньої потужності сигналу потрібно знати функцію щільності ймовірності.

Нерівномірне (нелінійне) квантування

Лінійні пристрої квантування легко реалізувати і легко зрозуміти - в цьому їх очевидне гідність. Разом з тим, вибір параметрів пристроїв рівномірного квантування не припускав ніяких знань про статистику амплітуд і кореляційних властивостях вхідного сигналу.

Нелінійні пристрої квантування, що забезпечують нерівномірний квантування, застосовуються тоді, коли виникає бажання врахувати статистику амплітуд і кореляційні властивості вхідного сигналу.

Існують додатки, для яких рівномірні пристрої квантування є найкращими. Це - обробка музичних сигналів, обробка зображень, контроль процесів і ряд інших. Для деяких інших додатків більш прийнятні нерівномірні Квант пристрої. Найважливішим прикладом такого роду є обробка мовних сигналів в системах зв'язку.

Людська мова характеризується унікальними статистичними властивостями.

Нерівномірне квантування може забезпечити краще квантування слабких сигналів і грубе квантування сильних сигналів. Значить, в цьому випадку шум квантування може бути пропорційним сигналу. Результатом є підвищення загального ставлення сигнал / шум - зменшення шуму для домінуючих слабких сигналів за рахунок підвищення шуму для рідко зустрічаються сильних сигналів. На рис. 1.9 порівнюється квантування слабого і сильного сигналів при рівномірному і нерівномірному квантуванні. Ступінчасті сигнали являють собою апроксимації аналогових сигналів (після введення спотворення внаслідок квантування).

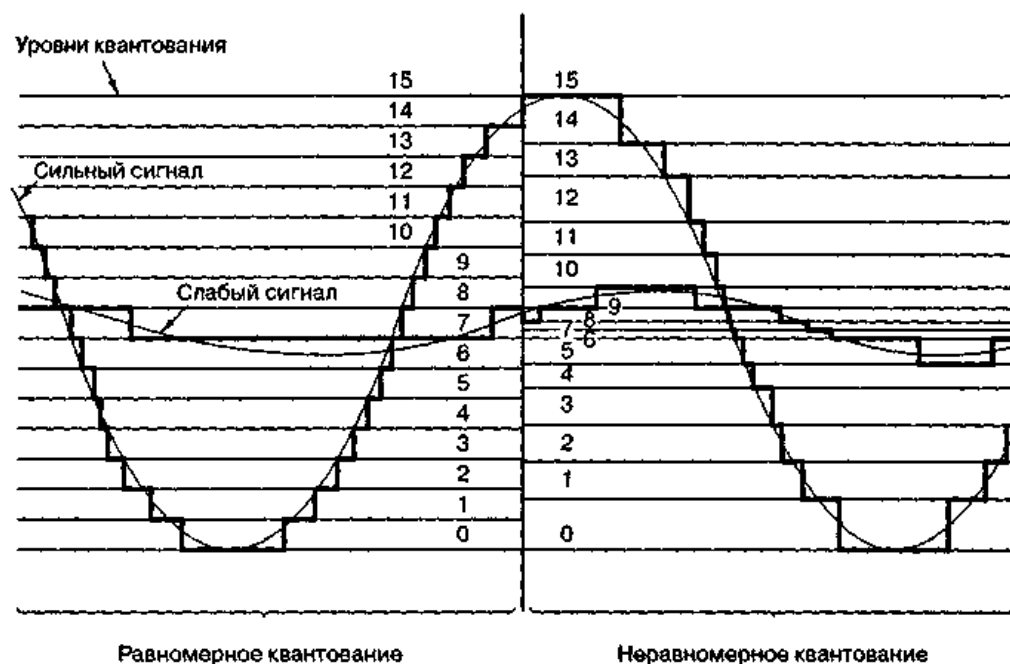


Рис.1.9. Рівномірне та нерівномірне квантування сигналів

Одним із способів отримання нерівномірного квантування є використання пристрою з нерівномірним квантуванням з характеристикою, показаної на рис. 1.10.а. Набагато частіше нерівномірне квантування реалізується в такий спосіб: спочатку вихідний сигнал деформується за допомогою пристрою, що має логарифмічну характеристику стиснення, показану на рис. 1.10.б, а потім використовується пристрій квантування з рівномірним кроком. Для сигналів

малої амплітуди характеристика стиснення має більш крутий фронт, ніж для сигналів великої амплітуди. Отже, зміна даного сигналу при малих амплітудах торкнеться більше число рівномірно розміщених рівнів квантування, ніж той же зміна при великих амплітудах. Характеристика стиснення ефективно змінює розподіл амплітуд вхідного сигналу, так що на виході системи стиснення вже не існує переваги сигналів малих амплітуд. Після стиснення деформований сигнал подається на вхід рівномірного (лінійного) пристрою квантування з характеристикою, показаної на рис. 1.10.в. Після прийому сигнал пропускається через пристрій з характеристикою, зворотної до показаної на рис. 1.10.б і званої розширенням, так що загальна передача не є деформованою. Описана пара етапів обробки сигналу (стиснення і розширення) в сукупності зазвичай іменується компандуванням.

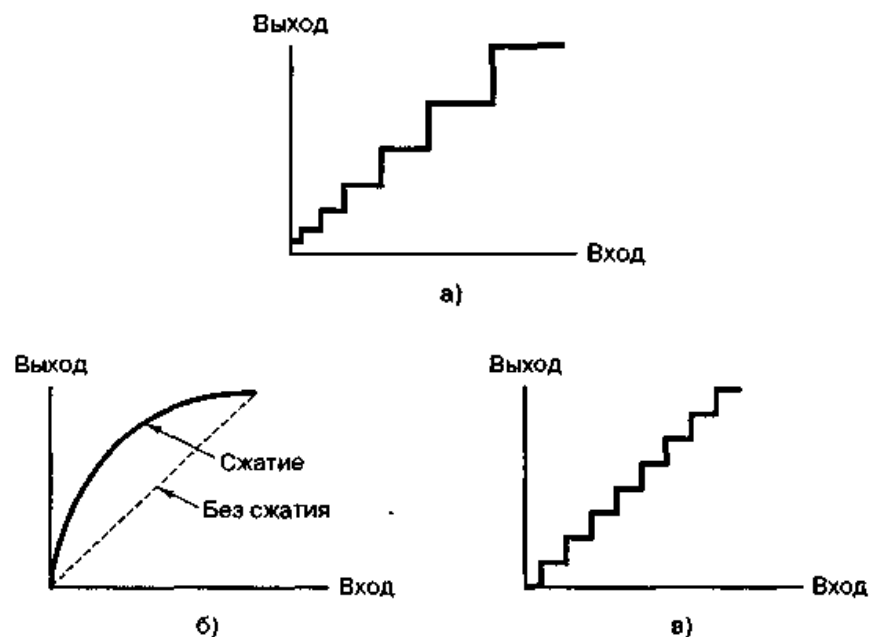


Рис.1.10. Приклади характеристик: а) характеристика нерівномірного пристрою квантування; б) характеристика стиснення; в) характеристика рівномірного пристрою квантування

Закон кодування

Неоднорідний процес квантування слідує певній ознаці, що називається законом кодування. Існує два типи законів кодування: безперервне та сегментоване.

У законах безперервного кодування інтервали квантування мають різну ширину, зростаючи від малих значень, що відповідають сигналам низького рівня, до більших значень, відповідних сигналам високого рівня.

У законах сегментованого кодування діапазон операцій ділиться на кінцеву кількість груп. Кожен інтервал однієї групи має однакову ширину, відрізняючись від інших груп.

Зазвичай використовувані закони кодування сегментовані.

G.711 А-закон (a-law) та μ-закон (u-law)

Два основні закони кодування, що застосовуються нині, є А-закон (a-law) та Мю-закон (u-law), які також відомі як g .711 кодек. А-закон (a-law) використовується в основному в європейських системах РСМ, а закон μ (u-law) використовується в американських системах РСМ.

А-закон утворений 13 прямими відрізками (насправді вони 16 сегментів, але три центральні сегменти вирівняні, тому їх зменшено до 13)

Математична постановка А-закону дорівнює:

$$y = \frac{Ax}{1 + \ln A}, \text{ для } x \leq \frac{1}{A}$$

$$y = \frac{1 + \ln(Ax)}{1 + \ln A}, \text{ для } \frac{1}{A} \leq x \leq 1$$

де $A = 87.6$ (1.5)

Параметр А приймає значення 87,6. X і Y являють собою вхідний і вихідний сигнал компресора

Математична постановка Мю-закону:

$$y = \frac{\ln(1 + \mu|x|)}{\ln(1 + \mu)}, \text{ де } \mu = 255$$

(1.6)

На наступному зображенні (1.11) графічно представлений А-закон:

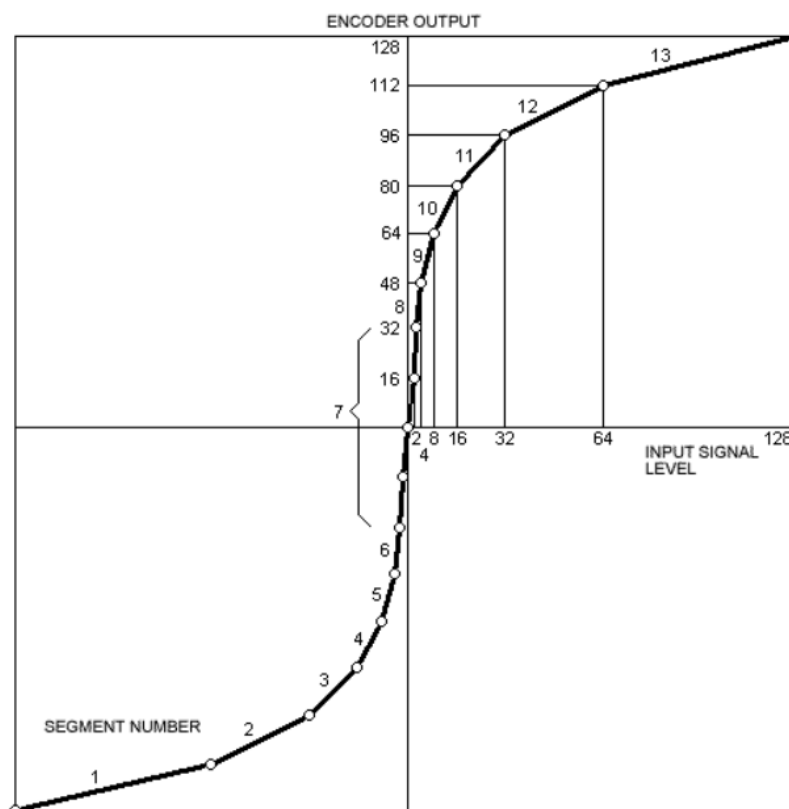


Рис.1.11. Графічне зображення А-закону

Диференціальна квантування (Диференціальна РСМ)

У голосових сигналах аудіалізація LF (низька частота), як правило, частіше зустрічається. З цієї причини рівень двох послідовних зразків зазвичай відрізняється дуже невеликою кількістю. Скориставшись цією обставиною, було створено диференційоване квантування. В диференційному квантуванні замість того, щоб кожен зразок обробляти окремо, його кількісно оцінюють та кодифікують різницю між вибіркою та попередньою. Оскільки кількість інтервалів квантування, необхідних для кількісної оцінки різниці між двома послідовними вибірками, менше необхідної для кількісного визначення одного ізольованого зразка, то диференційне квантування дозволяє зменшити частоту передачі, оскільки це пропорційно інтервалам квантування.

Дельта диференційованого квантування та ADPCM (Адаптивна дельта РСМ)

Якщо ми збільшимо частоту вибірки при диференційованому квантуванні, два послідовних зразки мають дуже невеликі відмінності у своєму

рівні. Тому для кількісної оцінки різниці можна використовувати один інтервал кількісної оцінки.

Для цього методу потрібен лише один біт для вибірки, і швидкість передачі (швидкість передачі бітів) буде дорівнює швидкості вибірки. Цей тип квантування відомий як дельта квантування.

У цій дельта-квантуванні рівень варіації виходу є унікальним. В інших типах дельта квантування варіація не фіксована і залежить від варіацій вхідного сигналу. Наприклад, ADPCM або Adaptive delta PCM засновані на динамічному підході шкали квантування в залежності від малих або великих відмінностей вхідного сигналу.

Кодифікація – розшифрування.

Кодифікація - це процес, за допомогою якого кількісно виражений зразок представляється двійковим числом з "1's" і "0's".

Зазвичай у телефонії 256 інтервалів квантування використовуються для представлення всіх можливих значень вибірки (наприклад, для G.711 або А закону та μ закону). Тому потрібно 8 біт для представлення всіх інтервалів ($2^8 = 256$). Інші кодеки, які використовують ADPCM або дельта квантування, використовують менші інтервали, і тому потрібно менше бітів для кодифікації зразків.

Пристрій, що здійснює квантування та кодифікацію, називається кодером.

Розшифровка - це процес, за допомогою якого реконструюються зразки, з числового сигналу. Цей процес проводиться в пристрої, який називається декодером.

Група кодера і декодера в одному і тому ж обладнанні називається кодеком.

ВАЖЛИВО: Якщо ми хочемо обчислити швидкість передачі кодеків, нам потрібно лише помножити частоту дискретизації, виражену в вибірках на секунду або Герцах на біти, необхідні для кількісного визначення кожного зразка, і це дає нам швидкість передачі бітів кодека.

У будь-якому випадку, оскільки існують складні кодеки зі стисненням, швидкість передачі бітів не завжди може бути виведена таким чином.

1.5. Моделювання квантування в Matlab

М-функція квантування в Matlab має вигляд:

[ind, xQ, D] = quantiz(x, partition, codebook),

де **x** - вектор відліків сигналу; **partition** - вектор кордонів інтервалів P_1, P_2, \dots, P_{L-1} ; **codebook** - вектор C_1, C_2, \dots, C_L значень кодової книги; **ind** - вектор-стовпець індексів (номерів) інтервалів; **xQ** - вектор-рядок квантованих значень сигналу; **D** - середній квадрат помилки квантування:

$$D = \frac{1}{N} \sum_{n=1}^{N_t} (x_n - \tilde{x}_n)^2 \quad (1.7)$$

Приклад застосування m-функції:

```
% квантование гармонического (50 Гц) сигнала единичной амплитуды
% частота дискретизации 1000 Гц
i=1:20;
x=cos(2*pi*50*i/1000);
partition=[-.75 -.25 .25 .75];
codebook=[-.825 -.5 0 .5 .825];
[ind, xQ, D] = quantiz(x, partition, codebook);
plot(i,x), hold on, stem(i,xQ)
legend('Исходный сигнал','Результат квантован.')
for n=1:4
prt(n,i)=partition(n); plot(i,prt,'k')
end
for m=1:5
cdbk(m,i)=codebook(m); plot(i,cdbk,'r-.')
end
figure
stem(i,ind), title('Индексы')
ind, xQ, D
```

Вихідний і квантований сигнали наведені на рис. 1.12, графік значень індексів наведено на рис. 1.13.

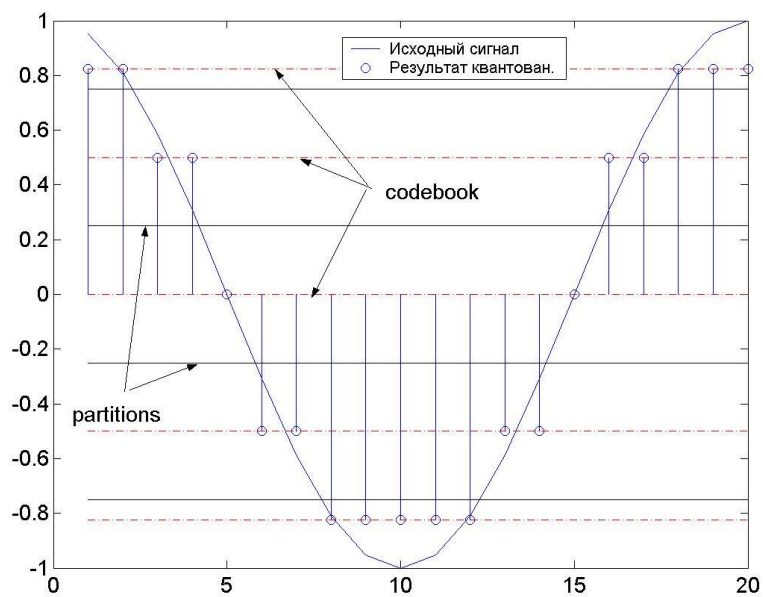


Рис.1.12. Вихідний і квантований сигнали

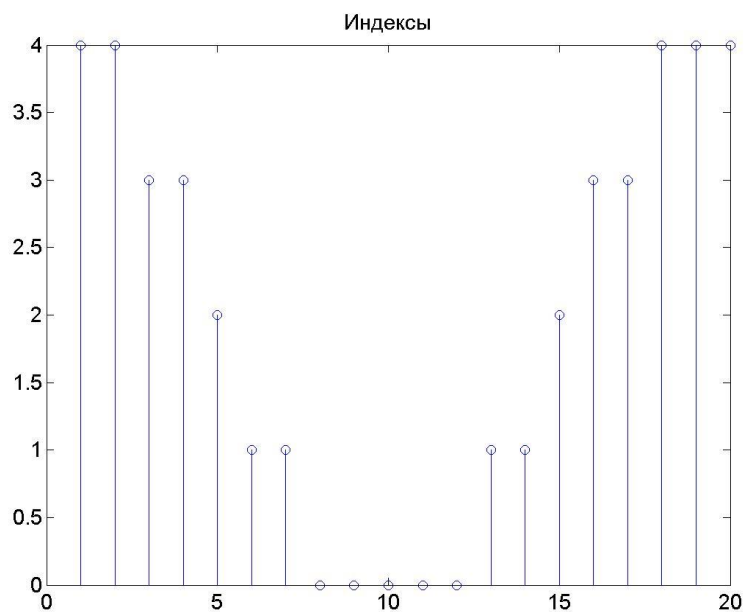


Рис.1.13. Графік значень індексів

1.6. Кодек GSC: мовний кодек GSM

Мовний кодек GSM - це конструктивна комбінація кодера і декодера, яка використовується як високо інтегрована телекомунікаційна складова. Метод кодування розроблений таким чином, що структура декодера залишається

рівномірною, незважаючи на різні логічні канали. Отримані кадри декодуються за допомогою методу оцінки максимальної ймовірності послідовності MLSE для досягнення оптимальної швидкості виявлення помилок та виправлення помилок. Для MS, еквалайзер та перетворювачі A/D і D/A також інтегровані в кодек. Голосовий транскодер підключений до кодека GSM як компонент VLSI. Це стискає оцифрований голосовий сигнал у потоці даних 13 кбіт/с. Функції мовного процесу, такі як VAD, шум комфорту та заміна кадру мови, також присвоюються кодеку.

Статистика показує, що продуктивність кодеку хороша, коли коефіцієнт переносника до перешкод вище 10 дБ.

GSM адаптивний багатошвидкісний кодек (AMR)

Кодек GSM AMR (Adaptive Multirate) описує детальне відображення з вхідних блоків 160 вибірок мови в 13-бітної уніфікованою формою ІКМ в кодовані блоки з 95, 103, 118, 134, 148, 159, 204 і 244 бітів і з кодованих блоків з 95, 103, 118, 134, 148, 159, 204 і 244 біта для виведення блоків з 160 відновлених мовних вибірок. Частота дискретизації становить 8000 вибірок / с, що призводить до швидкості передачі для закодованого потоку бітів 4,75, 5,15, 5,90, 6,70, 7,40, 7,95, 10,2 або 12,2 кбіт/с.

Схема кодування для режимів багатошвидкісного кодування використовує кодер лінійного прогнозування з порушенням по алгебраїчному коду (ACELP); багатошвидкісний ACELP-кодер називається MR-ACELP. Процедура транскодування, зазначена в стандарті GSM 06.90, застосовна для адаптивних багатошвидкісних каналів з повною швидкістю і половинної швидкості мовного трафіку (TCH) в системі GSM.

GSM AMR-кодер приймає свій вхід у вигляді 13-бітного уніфікованого сигналу PCM або від звукової частини мобільної станції, або на стороні мережі, від PSTN через 8-бітний А-закон або μ -закон до 13-бітного уніфікованому PCM

перетворення. Кодована мова на виході кодера доставляється в блок каналного кодера. У напрямку прийому виконуються зворотні операції.

- Широкосмуговий кодек AMR використовує вісім вихідних кодеків зі швидкостями 12,2, 10,2, 7,95, 7,40, 6,70, 5,90, 5,15 і 4,75 кбіт/с. Кодек заснований на моделі кодування з лінійним передбаченням з кодовим збудженням (CELP). Використовується лінійний прогноз (LP) 10-го порядку або короткостроковий синтезує фільтр. Фільтр синтезу основного тону реалізований з використанням так званого підходу адаптивної кодової книги.
- У моделі синтезу мови CELP сигнал збудження на вході фільтра короткочасного синтезу LP будується шляхом додавання двох векторів збудження з адаптивних і фіксованих (інноваційних) кодових книг. Мова синтезується шляхом подачі двох правильно обраних векторів з цих кодових книг через фільтр короткострокового синтезу. Оптимальна послідовність збудження в кодової книзі вибирається з використанням процедури пошуку «аналіз за синтезом», в якій помилка між вихідної і синтезованої промовою мінімізується відповідно до зваженої мірою спотворення. Ваговий фільтр використовує неквантовані параметри LP.
- Кодер працює з мовними кадрами тривалістю 20 мс, відповідними 160 вибірках, з частотою дискретизації 8000 вибірок / с. На кожні 160 вибірок мовної сигнал аналізується для отримання параметрів моделі CELP (коефіцієнти фільтра LP, індекси і коефіцієнти адаптивної та фіксованого кодових книг). Ці параметри кодуються і передаються. У декодері ці параметри потім декодуються, і мова синтезується шляхом фільтрації відновленого сигналу збудження через фільтр синтезу LP.
- Аналіз LP виконується двічі за кадр для режиму 12,2 кбіт/с і один раз для інших режимів. Для режиму 12,2 кбіт/с два набори параметрів LP перетворюються в пари лінійного спектра (LSP) і спільно квантуються з використанням квантування з поділом матриць (SMQ) з 38 бітами. Для інших режимів єдиний набір параметрів LP перетворюється в пари спектра

ліній (LSP) і квантів вектор з використанням векторного квантування з поділом (SVQ).

- Мовний кадр ділиться на 4 субкадрах по 5 мс кожен (40 вибірок). Параметри адаптивної та фіксованого кодової книги передаються в кожному субкадрі. Квантовані і неквантовані параметри LP або їх інтерпольовані версії використовуються в залежності від підкадрів. Затримка основного тону в розімкнутому контурі оцінюється в кожному іншому субкадрі (за винятком режимів 5,15 і 4,75 кбіт/с, для яких це робиться один раз за кадр) на основі зваженого по сприйняттю голосового сигналу.

Потім такі операції повторюються для кожного підкадрів.

GSM AMR декодер

У декодері на основі обраного режиму передані індекси витягуються з прийнятого потоку бітів. Індокси декодуються для отримання параметрів кодера в кожному кадрі передачі. Цими параметрами є вектори LSP, запізнювання основного тону, інноваційні кодові вектори, а також висота тону і інноваційні посилення. Вектори LSP перетворюються в коефіцієнти фільтра LP і інтерполують для отримання фільтрів LP в кожному субкадрі. Потім на кожному 40-елементному субкадрі:

Порушення створюється шляхом додавання адаптивних та інноваційних кодових векторів, масштабованих по їх відповідними коефіцієнтами посилення.

Мова відновлюється шляхом фільтрації збудження через фільтр синтезу LP.

Нарешті, відновлений голосовий сигнал пропускається через адаптивний фільтр поста.

Особливості AMR (Adaptive Multirate):

- Частота дискретизації 8 кГц / 13 біт (160 вибірок для кадрів 20 мс), відфільтрована до 200-3400 Гц.
- Кодек AMR використовує вісім вихідних кодеків зі швидкостями 12,2, 10,2, 7,95, 7,40, 6,70, 5,90, 5,15 і 4,75 кбіт / с.
- Генерує довжину кадру 95, 103, 118, 134, 148, 159, 204 або 244 біт для швидкостей AMR FR 4,75, 5,15, 5,90, 6,70, 7,40, 7,95, 10,2 або 12,2 кбіт / с відповідно. Довжини кадрів AMR HR різні.
- AMR використовує переривчасту передачу (DTX) з виявленням голосової активності (VAD) і генерацією комфортного шуму (CNG), щоб зменшити використання смуги пропускання під час періодів мовчання
- Алгоритмічна затримка становить 20 мс на кадр. Для швидкостей передачі 12,2 не існує «алгоритму» затримки перегляду. Для інших швидкостей затримка прогнозування складає 5 мс. Зверніть увагу, що є «фіктивна» затримка перегляду в 5 мс, щоб забезпечити плавне перемикавання кадрового режиму з іншими швидкостями.
- AMR є гібридним мовним кодером і, як такий, передає як мовні параметри, так і сигнал форми сигналу: кодування з лінійним передбаченням (LPC) використовується для синтезу мови з залишкового сигналу. Параметри LPC кодуються як лінійні спектральні пари (LSP); залишкова форма сигналу кодується з використанням лінійного прогнозування з порушенням по алгебраїчному коду (ACELP).
- Складність алгоритму оцінюється в 5 з використанням відносної шкали, де G.711 дорівнює 1, а G.729a - 15.
- Тестування PSQM в ідеальних умовах дає середні оцінки 4,14 для AMR (12,2 кбіт / с) в порівнянні з 4,45 для G.711 (μ-закон)
- Тестування PSQM в умовах стресу в мережі дає середні оцінки 3,79 для AMR (12,2 кбіт/с) в порівнянні з 4,13 для G.711 (μ-закон)

ВИСНОВКИ до розділу 1

В даному розділі ми розглядали кодеки, а саме їх класифікацію. За допомогою кодеків без втрат вся аудіо- чи відеоінформація, що міститься в оригінальних даних, зберігається у початковій якості. Кодеки, які призводять до втрати якості, можуть забезпечити більше стиснення, щоб використовувати меншу пропускну здатність під час передачі.

Коли ми здійснюємо голосові дзвінки через Інтернет або в інших цифрових мережах, голос повинен бути закодований в цифрові дані, і навпаки. У цьому ж процесі дані стискаються таким чином, що передача відбувається швидше, а досвід виклику - кращий. Це кодування досягається кодеками.

AMR забезпечує якість голосових зв'язків, що наближається до швидкості 7,4 кбіт/с, з якістю, близькою до платної, при більш низьких швидкостях і більшій стійкості та вдосконаленням відтворенням неречових звуків з більш високими швидкостями. Належно AMR є найновішим з 8 швидких, сьогодні він може вважатися найбільш широко використовуваним кодеком в світі, працюючи на 2 мільярдах безпроводних телефонах, а також на базових станціях та інших мережевих обладнаннях.

Незважаючи на те, що існує безліч голосових кодеків, найбільш популярними є два, G.711 і G.229. G.711 забезпечує нестислий високоякісний голос. G.729, з іншого боку, стискається, так що він використовує меншу пропускну здатність на шкоду якості.

РОЗДІЛ 2

ЯКІСТЬ КОДОВАНИХ СИГНАЛІВ

Як показано у першому розділі роботи, кодеків існує багато, і кожен з них придатний для певних цілей. Необхідні кодеки для передачі цифрових сигналів. Цифрові сигнали подаються у вигляді чисел, розділених за часом та амплітудою. Розділення за часом – процес дискретизації. За амплітудою – квантування. Усі кодеки в основному виконують ці дві функції.

З дискретизацією усе зрозуміло: чим більше відліків сигналу зберігається у одиниці часу, тим якісніший сигнал. Квантування – цікавіший процес: необхідно знаходити баланс між розміром сигналу, тобто швидкістю його передачі, та якістю сигналу.

Розглянемо на прикладі кодеку G.711. Стандартно, подаючи на його вхід 16-бітний сигнал, на виході отримаємо сигнал з рівнем квантування 8 біт. Як було показано раніше, в його роботі виконується нерівномірне квантування, яке передбачає використання логарифмічної шкали, замість лінійної, і змодельовавши цей кодек, можна керувати його параметрами, регулюючи рівень та глибину квантування.

Існують застосунки, для яких рівномірні пристрої квантування є найкращими. Серед них – обробка музичних сигналів, зображень, контроль процесів. Обробка мовленнєвих сигналів є найважливішим прикладом застосування саме нерівномірного квантування. Пояснюється це унікальними статистичними властивостями людської мови. Одна з таких властивостей показана на рис. 2.1.

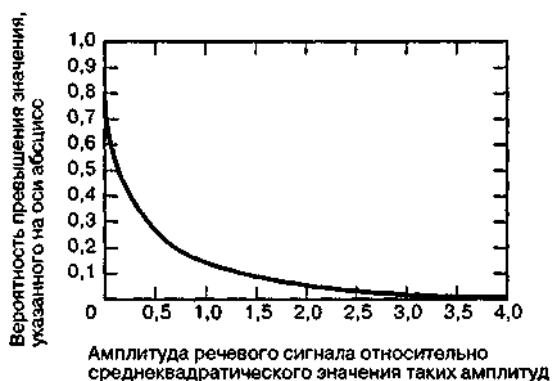


Рис. 2.1. Статистичний розподіл амплітуд мови.

На осі абсцис відкладені амплітуди мовленнєвого сигналу, нормовані за середньоквадратичним значенням амплітуд, по осі ординат – вірогідність. Для більшості каналів домінують дуже низькі рівні сигналів: 50% часу напруга, що характеризує енергію мови, складає менше чверті середньоквадратичного значення. Значення з великими амплітудами зустрічаються відносно нечасто: тільки 15% часу напруга перевищує середньоквадратичне значення.

При передачі мови система з рівномірним квантуванням буде неекономною – багато відліків квантування будуть використовуватись доволі рідко. Крім того, у такій системі шум квантування буде однаковим для всіх амплітуд сигналу. У наслідку, при такому квантуванні, відношення сигнал/шум буде гірше для сигналів низьких рівнів, ніж для сигналів високих рівнів.

Для моделювання кодеку G.711, було створено програму (Дод. А), в якій порівнюються сигнали, отримані шляхом лінійного квантування та шляхом квантування за допомогою кодеку. На рис. 2.2 та рис. 2.3 зображено графіки отримані у результаті виконання програми.

Якщо порівнювати графіки, можна побачити різницю між двома типами квантування, яку дуже добре видно на високих амплітудах: при нерівномірному квантуванні різниця між сусідніми піками більш помітна, ніж при лінійному, де спад амплітуди плавніший.

Порівнюючи результат на слух, можна сказати що при лінійному квантуванні помітний шум, трескіт. При нелінійному його майже немає, до того ж якість сигналу помітно вища.

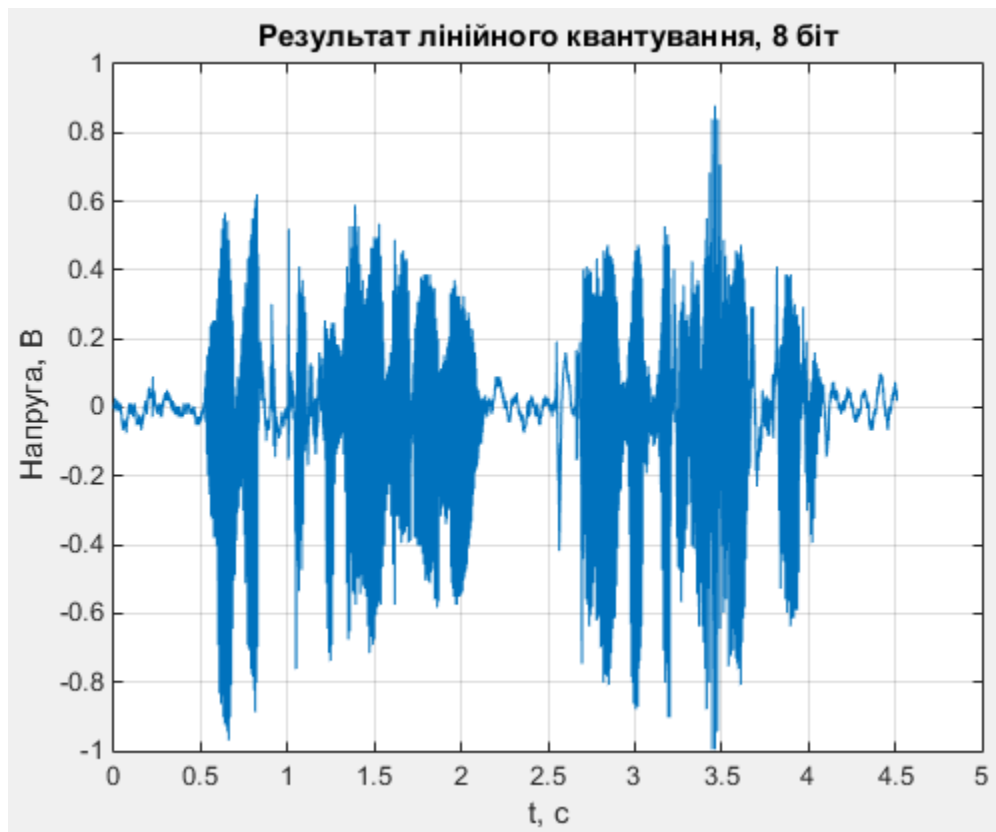


Рис. 2.2. Результат лінійного квантування, 8 біт.



Рис. 2.3. Результат нелінійного квантування, 8 біт.

Більш помітною різниця стає при зменшенні рівнів квантування сигналів. Так, зменшивши рівень до 6 біт, отримаємо результати з рис. 2.4 та рис. 2.5.



Рис. 2.4. Результат лінійного квантування, 6 біт.



Рис. 2.5. Результат нелінійного квантування, 6 біт.

Все вище сказане про 8-бітну модель квантування можна сказати і про 6-біт. Але зараз вже варто порівняти графіки з рис. 2.4, рис. 2.5 та рис. 2.6 (чистий сигнал, лінійне квантування 6 біт та нелінійне квантування 6 біт). Оскільки основна частина сигналу міститься в інтервалах малих амплітуд, варто звернути увагу на них, а саме на те, що при нерівномірному квантуванні сигнал більш схожий на чистий, ніж при лінійному. Про звук також можна сказати що помітно якісніший сигнал, з меншим трескітом саме при нелінійному квантуванні.

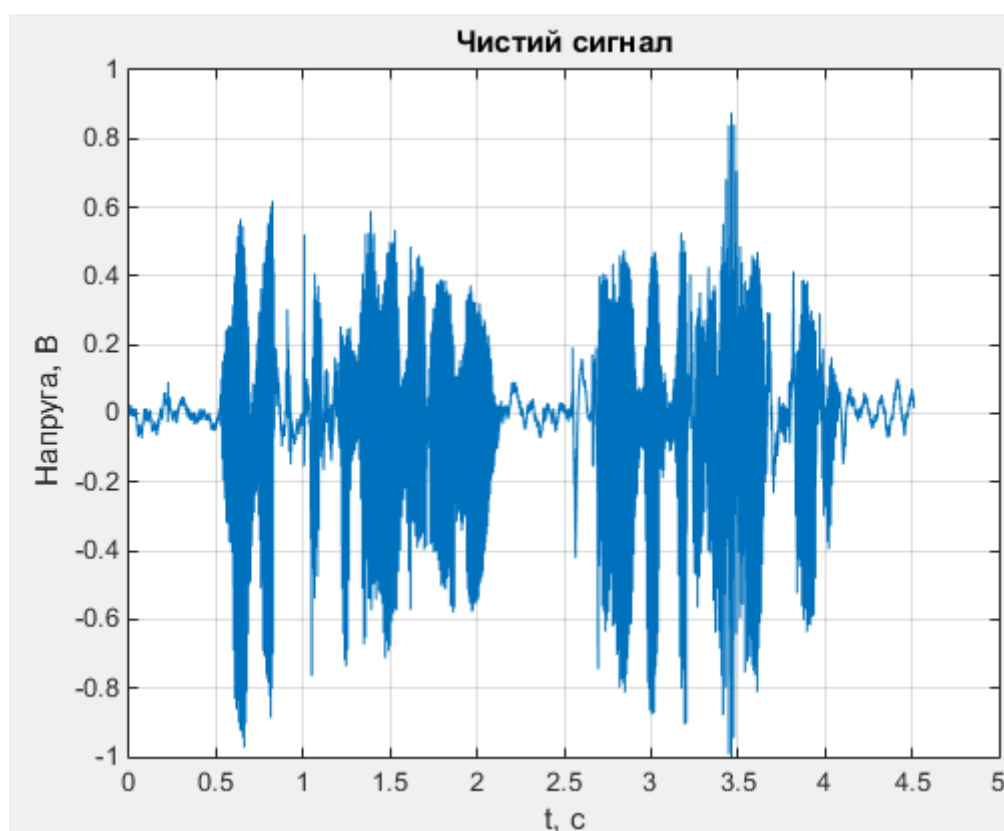


Рис. 2.6. Чистий сигнал.

ВИСНОВКИ до розділу 2

Було виконано моделювання кодека G.711. Його особливість полягає у використанні нерівномірного квантування замість лінійного. Порівнявши графіки та звучання сигналів, можна сказати що при нерівномірному квантуванні сигнал мав менше шумів та загалом звучав краще ніж лінійно-квантований.

Цей кодек був запропонований у 1972 році саме для телефонних ліній, а тому на першому місці стає економія ресурсів та часу. На прикладі моделі видно що замість лінійно-квантованого 8-бітного сигналу, можна використати нерівномірне, 6-бітне квантування, при цьому втративши у якості, але зменшивши ресурс передачі з 64 біт/с до 48 біт/с.

За допомогою кодека можна забезпечити вищу швидкість передачі сигналу, при цьому слабо вплинувши на його якість, що показано на прикладі моделювання G.711.

ВИСНОВКИ

В даній роботі виконано аналіз використання кодеків в цифрових лініях зв'язку, а також експериментально оцінено якість сигналів при застосуванні кодеку G.711.

Виконано класифікацію кодеків. Показано, що існує два типи кодеків – це програмні та програмно-апаратні. Детально ознайомились з мовними, мобільними та VoIP кодеками. Для телефонії кодеки працюють за принципом вибірки, квантування та кодування. Середня оцінка думки (MOS) допомагає нам суб'єктивно оцінити якість звуку, який надходить до слухача після всього процесу стиснення, передачі та декомпресії. Процес обробки сигналу для кодека G.711 виконується в три етапи: вибірка, квантування та кодифікація. Кодек G.711 забезпечує нестислий голос високої якості. Квантування поділяється на рівномірне (лінійне) та нерівномірне (нелінійне). Критерієм якості пристрою квантування є дисперсія похибки квантування. Нині застосовуються два основні закони кодування – А-закон та Мю-закон. GSM AMR забезпечує якість зв'язку, близькою до платної, при більш низьких швидкостях та із більшою стійкістю.

Виконано моделювання кодеку G.711 в Matlab. Наведено результати експериментальних досліджень, з яких виходить, що при нелінійному квантуванні сигнал має менше шумів та звучить краще лінійно-квантованого. Експериментально показано, що кодек G.711 дає змогу передавати мовний сигнал із достатньо великою швидкістю та із малим спотворенням навіть при зниженні глибини квантування до 6 бітів.

ПЕРЕЛІК ПОСИЛАНЬ

1. Sudeept Bhatnagar, Samrat Ganguly / VoIP: Wireless, P2P and New Enterprise Voice Over IP / 2008
2. General aspects of digital transmission systems [Електронний ресурс]. – Режим доступу: https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-G.711-198811-I!!PDF-E&type=items
3. Продеус А.Н., Родінова М.В. Безпаперова технологія проведення практикумів із статистичної обробки сигналів. – Електроніка и связь, №20, 2003, pp.117-120
4. Tamal Chakraborty, Iti Saha Misra, Ramjee Prasad / VoIP Technology: Applications and Challenges / 2019
5. Гульяев А. Имитационное моделирование в среде Windows. – С-Пб, КОРОНА принт, 1999. – 287 с.
6. Скляр Б. Цифровая связь. Теоретические основы и практическое применение. – М., С-Пб, К., изд. дом «Вильямс», 2003. – 1092 с.
7. Калюжний О.Я. Моделювання систем передачі сигналів в обчислювальному середовищі MATLAB-Simulink. – К., «Політехніка», 2004. – 135 с.
8. Codecs and VoIP protocols [Електронний ресурс]. – Режим доступу: <http://www.en.voipforo.com/codec/codecs.php>
9. Дидковский В.С., Дидковская М.В., Продеус А.Н. Акустическая экспертиза каналов речевой коммуникации. Монография. – К., «Имекс-ЛТД», 2008. – 420 с.
10. Гольдштейн Б.С., Пинчук А.В., Суховицкий А.Л. / IP ТЕЛЕФОНИЯ / МОСКВА «РАДИО И СВЯЗЬ» / 2001
11. VoIP Mechanic [Електронний ресурс]. – Режим доступу: <https://www.voipmechanic.com/what-is-voip.htm>
12. Jacek Biala / Mobilfunk und Intelligente Netze: Grundlagen und Realisierung mobiler Kommunikation / 2013

13. The Adaptive Multirate Wideband speech codec (AMR-WB)
[Электронный ресурс]. – Режим доступа:
https://www.researchgate.net/publication/3333904_The_Adaptive_Multirate_Wideband_speech_codec_AMR-WB

Додаток А. Програма моделювання кодеку G.711 в середовищі Matlab

```
%G.711 A-law encode
% демо кодування-декодування із варіюванням кількістю рівнів
квантування
% вхідні дані:
clear; clc
name = input('Введіть назву файлу:      ', 's');
[x, fs] = audioread(name);
x = x/max(abs(x));           % нормировка сигнала по максимуму
disp('Слухаємо вхідний сигнал')
disp(' ');
soundsc(x, fs);
pause(length(x)/fs);
t = 0:1/fs:(length(x)-1)/fs;

%% нелінійне квантування
% компресія перед квантуванням
disp('A = 1 - лінійне квантування');
disp('A = 86.7 - типове нелінійне квантування');
disp(' ');
A = input('Обераємо ступінь нелінійності A (від 1 до 86.7):
');
y = compand(x, A, max(x), 'A/compressor');

% квантування
b = input('Кількість бітів квантування (від 1 до 16):      ');
L = 2^b;                               % кількість рівнів квантування
q = 2/L;                               % розмір одного кванта
codebook = -1+q/2:q:1-q/2;             % середини квантів (кодова
книга)
partition = -1+q:q:1-q;                 % границі між
квантами
[~, yquan, ~] = quantiz(y, partition, codebook); % квантований
сигнал
pause(5)

% декомпресія після квантування
xquan = compand(yquan, A, max(yquan), 'A/expander');
disp('Слухаємо вихідний сигнал')
soundsc(xquan, fs);
%% побудова графіків
plot(t, x); grid;
% title('Чистий сигнал'); xlabel('t, c'); ylabel('Напруга, В');
```